

矩阵的几何意义

Guangyao Zhao

2022-08-04

Contents

矩阵的概念	1
矩阵加法的意义	2
矩阵与向量的乘法的意义	2
矩阵与矩阵的乘法	2
矩阵与线性变换的几何意义	3
矩阵乘法运算律的几何意义	3
矩阵秩的几何意义	4
矩阵特征值和特征向量的几何	4
矩阵的等价、相似与合同	4

矩阵是整个线性代数的重中之重，占据了整个线性代数的一半内容。矩阵刻画的是向量之间的变换关系，即将向量 a 经过矩阵 A 转换为向量 b 。变换后的向量可能仍在原向量空间，也可能被伸缩。

矩阵的概念

1. 矩阵可以是一个简单的统计表。比如现在有两个矩阵。矩阵 1 和 2 的第一列均为产品名称，之后的每一列对应月销量。那么两个矩阵相加的意义就代表两年的各个月份产品销量之和。矩阵作为统计鼠标的概念比较简单。
2. 还有一种是在机器学习广泛采用的形式，数表代表着样本。列为样本，行为样本特征。
3. 数表的矩阵是静态概念，无法深入展开，必须进入函数系数到线性映射的概念。即矩阵是向量和向量之间的线性对应关系。这一概念非常重要。

矩阵加法意义

矩阵的加法比较直观，即为对应位置数值相加即可：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \\ c_{11} & c_{12} & c_{13} \end{bmatrix} + \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ b_{21} & b_{22} & b_{23} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} + \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ b_{31} & b_{32} & b_{33} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

矩阵与向量的乘法意义

对比向量的内积，矩阵向量乘法稍显复杂，如果第一次见的话可能完全不知所以然。其实本质上来说就是一个向量组与一个向量的同时表示而已，本质上还是向量的内积。

矩阵与向量（矩阵）的乘法的意义深远，其代表着线性代数中的线性变化。线性变换又可以说是线性代数中核心中的核心。它描绘的是矩阵是如何在向量空间中，甚至不拘泥于当前向量空间到更高或更低的空间中的千变万化：

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} \end{bmatrix} \end{aligned}$$

矩阵与矩阵的乘法

从矩阵与矩阵的乘法是矩阵与向量的乘法的进一步推广，两个矩阵之间相乘的要求 $n \times m \times m \times k$ 可以看到矩阵乘法对两个矩阵的形状是有要求的，即矩阵 \mathbf{A} 的列要和矩阵 \mathbf{B} 的行数相等。其实本质就是多组向量间的内积：

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{bmatrix} \end{aligned}$$

任广千老师在书中用几何的方式形象的描绘了矩阵是两个坐标系之间的过渡或转换媒介，感兴趣的可以去阅读下原文。这里需要注意的是矩阵在转换坐标系的时候不仅可以在原向量空间中转换，而且是可以伸缩原向量空间。假设原向量空间时 3 维：

1. 若转换前后向量空间不变，则矩阵形状为方阵，即 3×3
2. 若转换前后向量空间空间缩小为 2 维，则矩阵形状为 2×3
3. 若转换前后向量空间空间扩大为 4 维，则矩阵形状为 4×3

在向量组一章有提到旋转矩阵的概念，即旋转前后坐标系仍旧单位正交。这也是一种对向量的变换。旋转矩阵的列向量表示逆时针旋转，行向量表示顺时针旋转。原则上大家都会使用列向量，所以我个人猜想这也是和象限的顺序有某种关系吧。

矩阵与线性变换的几何意义

矩阵的初等变换有三个，分别代表不同的意义：

1. 镜像：行交换， $T(i, j)$
2. 伸缩：某一行乘以一个非零系数 k ， $T(i(k))$
3. 切换：某一行乘以一个非零系数 k 加到另一行， $T(i, j(k))$

对于二阶单位矩阵 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 分别进行初等变化可得：

1. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
2. $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$ 或者 $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$
3. $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ 或者 $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

三个变换代表着三个不同的变换，本质上还是针对基向量平行四边形的变换。任何一个可逆矩阵都可分解为基本初等矩阵 (1), (2), (3) 的乘积。

矩阵乘法运算律的几何意义

1. 既然矩阵代表着对向量对基的线性变化，那么矩阵的乘法就可视为对其连续变换。
2. 矩阵不满足交换律，即先后线性变换的意义不同。
3. 矩阵不满足消去律，即 $AC = BC$ 不可想当然的简化为 $A = C$ 。具体得说，当 C 是奇异矩阵（非满秩）时不满足消去律，可逆矩阵（满秩）时满足。本质上类似于代数里的 $ac = bc$ ，当 $c = 0$ 时就不满足消去律了。

矩阵秩的几何意义

1. 我一直都特别喜欢线性代数中『秩』的翻译，秩就是秩序嘛。矩阵是一种线性变换，谈到线性变换又脱离不了线性空间，那么矩阵的秩就自然而然代表了矩阵的线性空间，张成二维还是三维空间，是该矩阵的一个固有属性，无论矩阵怎么进行初等变换，它的秩不变，还是维持矩阵内在的秩序。
2. 用数学术语描述秩：方阵里面，有几个行向量是多余的，就会有几个列向量是多余的。对于所有的矩阵，不想管的行的数等于不相关的列的数，就是秩数。
3. 那么矩阵的秩对于线性变换又意味着什么呢？比如现有矩阵 A 对 n 维向量 v 进行线性变换，无论 n 的值是多少，变换后的维数都为矩阵 A 的秩数。矩阵是将向量 v 映射到更高维的空间还是更低维的空间甚至说仍停留在原空间内，都是矩阵的秩数说了算。

矩阵特征值和特征向量的几何

1. 矩阵 A 和一个向量 a 相乘，如果和一个数 λ 和该向量相乘得到的结果一样，即 $Aa = \lambda a$ ，则称 a 是矩阵 A 的特征向量，其中 λ 是特征值。
2. 通俗地理解上面的表达就是，在矩阵对原线性空间变换时（牢牢记住，矩阵变换的对象永远都可以理解为对原线性空间基向量的变换，向量是依托于基向量的啊），有些向量变换前后方向改变，大小不变；有些向量变换前后方向改变大小不变；有些向量变换前后方向大小都会改变。对于上式子定义，矩阵 A 对某个向量 a 只进行了伸缩变化，伸缩率为 λ ，也就是说经过矩阵 A 线性变换后该向量的方向没变！再进一步就可以说 a 描述的就是那些线性变换前后方向不变的向量：方向大小都不变的伸缩率为 1，方向不变大小变的向量的伸缩率为 λ 。由于该向量太特殊了，和原空间中的其它向量比具有很强的特征，所以它便被称之为『特征向量』，伸缩率是描述特征向量的一个实数值，所以该伸缩率就被称之为『特征值』。
3. 实对称矩阵有一个非常好的性质，它的特征向量正交。
4. 相似矩阵的数学表达是 $A = PBP^{-1}$ ，先说结论，相似矩阵指的是两个矩阵有一样的线性变换，只是基向量不同而已。 P 和 P^{-1} 就是一对互逆运算。

矩阵的等价、相似与合同

1. 等价: $B = PAQ$
2. 相似: $B = P^{-1}AP$
3. 合同: $B = C^TAC$

可以注意到，条件一次比一次苛刻，根据 P 和 Q 的关系的不同，可以依次分类为等价、相似和合同，矩阵 A 和 B 的关系也一次比一次密切。换句话说讲 A 转换为 B 一次比一次容易。